PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 5/00, 1/20, 1/14, 1/16, 1/18, C12P 21/06, C07K 1/00 — A1

(11) International Publication Number: WO 97/01627

(43) International Publication Date: 16 January 1997 (16.01.97)

(21) International Application Number: PCT/US96/10858

(22) International Filing Date: 27 June 1996 (27.06.96)

(30) Priority Data:
60/000,822 27 June 1995 (27.06.95) US
08/553,006 3 November 1995 (03.11.95) US

(71) Applicant: IGEN, INC. [US/US]; 16020 Industrial Drive, Gaithersburg, MD 20877 (US).

(71)(72) Applicant and Inventor: YOST, Pamela, B. [US/US]; 19617 Selby Avenue, Poolesville, MD 20874 (US).

(72) Inventors: PILON, Aprile, L.; 11333 Rambling Road, Gaithersburg, MD 20879 (US). LOHNAS, Gerald, L.; 4 Slate Mills Court, Catonsville, MD 21228 (US). ROBERTS, Steven, F.; Apartment 20, 2216 Washington Avenue, Silver Spring, MD 20910 (US).

(74) Agents: RYAN, John W.; IGEN, Inc., 16020 Industrial Drive, Gaithersburg, MD 20877 (US) et al.

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: HIGH-LEVEL EXPRESSION AND EFFICIENT RECOVERY OF UBIQUITIN FUSION PROTEINS FROM ESCHERICHIA COLI

(57) Abstract

A process for obtaining high yields of recombinant protein using a ubiquitin-peptide fusion system. The system includes (1) a ubiquitin-specific protease, which cleaves only C-terminal extensions from ubiquitin, and (2) an expression system consisting of a suitable *E. coli* or other host strain paired with a plasmid that encodes the ubiquitin fusion. The process produces a variety of proteins and peptides including those ranging in size from 8 to 70 amino acids. The protein or peptide can be purified in a three step purification process which includes an enzymatic cleavage reaction and insures cost-effective production of gram quantities of peptides. The process typically generates over 2 grams of pure peptide from 12 grams of ubiquitin fusion, using the cell paste generated from a single 10-liter fermentation, and is readily scaled up. There is typically good correlation between increased culture volume, increased cell density and increased product yield.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

HIGH-LEVEL EXPRESSION AND EFFICIENT RECOVERY OF UBIQUITIN FUSION PROTEINS FROM ESCHERICHIA COLI

Field of Invention:

The invention generally relates to the field of molecular biology and more specifically to fusion proteins involving Ubiquitin and their use to scale up fermentation processes and enhance yields of recombinantly produced proteins.

Background:

Ubiquitin has been used as a fusion partner for the synthesis of recombinant peptides and proteins with other host-vector systems. Ubiquitin fusions expressed at 20% of the total cell protein, using the yeast ubiquitin coding sequence under control of the lambda repressor in *E. coli* strain AR58, have been reported (2). Peptide libraries have been synthesized as ubiquitin fusions at yields of 60 mg/liter of culture (20). Yields as high as 1 mg of peptide product per liter of bacterial culture have been reported by Yoo and Rechsteiner (7). However, there is no literature pertaining to the performance of ubiquitin fusion expression systems in fermentor cultures.

Ubiquitin, a 76 amino acid protein, is an integral part of the pathway that mediates protein degradation in eukaryotic cells. Ubiquitin is covalently attached to intracellular proteins that are targeted for rapid turnover. Branched chains of ubiquitin are then conjugated to the

first, allowing the target protein to be directed to and/or recognized by the proteosome, a multisubunit macromolecular structure that degrades the target protein to amino acids and peptides and recycles the ubiquitin monomers (1). Ubiquitin itself is extremely resistant to proteolytic degradation and has a long half-life (180 minutes) when expressed in *E. coli* (2). Several proteins have been expressed as C-terminal ubiquitin fusions, some of which cannot be produced at all in bacteria (3, 4) or yeast (5, 6) as independent entities. Even short peptides have been produced efficiently in *E. coli* as ubiquitin fusions (7). Ubiquitin is thought to stabilize the nascent peptide chain as it comes off the ribosome during translation, and to act as a chaperonin to fold the recombinant protein more efficiently (2).

Several issues remain problematic with *E. coli* expression including the avoidance of proteolysis and the production of proteins in a soluble or native state (8). In *E. coli*, recombinant proteins have been expressed at levels as high as 50% of the total cellular protein, particularly at low cell density (9). While this should result in tens of grams per liter in high cell density fermentations, production at this level has not been reported. This is partially due to a decrease in the metabolic efficiency of *E. coli* at high cell density, and to increased *in vivo* protein turnover. To ensure adequate expression, nutrients may be supplied during the induction phase of an extended fed-batch culture (10, 11). However, addition of nutrients may result in lower yields

because of nutrient interaction with various intracellular pathways and functions. For example, addition of excess glucose often results in increased organic acid production which decreases overall yield (e.g., acetate, (12)). High concentrations of TCA-cycle intermediates and amino acids have resulted in reduced protein yield due to the altered regulation of the TCA cycle, which must be fully operative for maximum expression (13).

Many reports have indicated that sensitivity to proteolytic attack is also responsible for low expression levels at high cell density. Most often, differences in stability are attributed to the specific protein (state of folding) and not to the conditions by which the proteins were expressed (14). Indeed, many protein-specific characteristics have been elucidated which lead to enhanced degradation. Yet, it is also well known that the activity of the attacking proteases or peptidases depends on cellular stress (15, 16). For example, Harcum and Bentley demonstrated that induced expression of chloramphenicol-acetyl-transferase (CAT) elicits several proteolytic activities that also appear during the stringent stress response (17, 18). Others have noted that overexpression of "abnormal" proteins amplifies stress protein levels and protease La (19). Since ubiquitin is an eukaryotic stress protein, it appears to be less sensitive to the elevated cellular stress that often appears in bacteria during overexpression conditions.

SUMMARY OF THE INVENTION

An object of this invention is to avoid or minimize proteolysis with *E. coli* and like expression systems and to enhance the production of proteins in a soluble or native state. A further object of the invention is the consistent high specific yield. A still further object is to obtain over one gram per liter yield of fusion protein when the fusion is soluble and the two-step purification applies.

The combination of optimal codons for individual peptide fusions are thought to be in part responsible for the high specific yields which makes this technology so valuable. Focussing on maximizing biomass in the cultures as this invention does, rather than manipulating the physiology of the organism, increases the specific yield, which increases overall yield by decreasing the amount of downstream processing.

The term fermentation, as used herein, is intended to refer to processes involving the production of recombinant protein products.

In the system of this invention, ubiquitin fusions are employed. They are extremely stable and are expressed at high levels in soluble form. *E. coli* has been used as a model for ubiquitin-peptide fusion systems in 10-liter batch cultures. Ubiquitin fusion proteins are produced at very high levels in *E. coli* and related hosts.

The specific yield is defined as the percentage, taken as a ratio, of recombinant protein product to total cellular protein, as measured by densitometry of SDS-PAGE gels run on whole cell samples lysed in loading buffer and loaded directly. Specific yields in the *E. coli* ubiquitin fusion system, grown and induced as described, exceed 50% and approach 90%. The highest reported accumulation of a recombinant protein in *E. coli* is 50% of the total cellular protein, i.e., 50% specific yield (9). Another *E. coli* expression system claims 40% specific yields for some recombinant proteins (29). Under the conditions specified in the invention, this level of expression exceeds 50% specific yields which results in yields of recovered fusion protein of over one gram of fusion protein per liter of bacterial culture. This protein is both soluble and recoverable. The majority of protein in supernatants of lysed cells is product. The fusion can be further purified from host proteins with an 85°C heat step, in which most of the host proteins precipitate while the ubiquitin fusion stays in solution.

The ubiquitin fusion process is outlined in Figure 1. Briefly, the process involves cloning the peptide or protein of choice as a C-terminal ubiquitin fusion into a suitable expression vector. Suitable peptides or proteins include, but are not limited to, a size range between 2 amino acids and 6000 amino acids, more typically between 6 amino acids and 500 amino acids, and specifically between 8 amino acids and 70 amino acids as shown by sequences specified in Table 4.

T7, trc, lac, and λpPL are preferred. Additional suitable promoters include lacUV5, tac, trc, trp, lexA, malE, araC, and nifA.

According to the invention, clones are screened for good expression and the correct DNA sequence using techniques which are well known to those of ordinary skill in the art. Fusion proteins are then expressed and isolated and the product is enzymatically released from ubiquitin by an ubiquitin-specific protease. The specificity of these proteases is exquisite so that product is not degraded. These proteases include, for example UCH-L1, UCH-L2, UCH-L3, YUH1, UBP1, UBP2, UBP3, UBP4, UBP5, UBP6 and fat facets (Drosophila). See also U.S. Patent No. 5,212,058 and International Application No. WO 91/17245 in this regard. Furthermore, any N-terminal amino acid on the product is possible because the cleavage site always follows the glycine-76 residue of ubiquitin. Once released, the product is purified and lyophilized.

The invention also includes a means of amplifying expression levels by coupling a chemical induction using, for example, isopropylthiogalactoside (IPTG) with heat shock. A 12°C heat shock, upon induction of the ubiquitin fusion with IPTG, produces an excellent yield of recoverable protein. Heat precipitation of host proteins is dependent upon the conditions under which the culture is grown and induced. A heat shock coupled with induction results in the stabilization of the host proteins, such that the benefits of the 85°C heat purification step are lost. However, under these conditions, the

ubiquitin fusion is overexpressed to greater than 90% of the total soluble protein, making the 85°C purification incubation unnecessary. The extraordinary stability of ubiquitin to heat and to proteases, coupled with the high levels of expression attainable in *E. coli*, make it an ideal fusion partner for a bioprocess.

In addition, the invention includes the development of a transformant strain of *E. coli* that overexpresses a recombinant protein in excess of 50% of the total protein, preferably at least 1.0 gram per liter. (See strain A37 as an example). Transformation of the host is effected by expression plasmids, such as pDSUb, pPX1, pRSET-Ub, pPL-Ub. Other suitable host cells can be employed, such as bacteria other than *E. coli*, yeast, insect, plant or mammalian cells.

Further, the invention includes processes for growth of bacterial and other host cells and induction of synthesis of a recombinant protein that yields the desired product in excess of 50% of the total protein or one gram per liter of bacterial culture. The induction includes chemical induction of synthesis of a recombinant protein, and if desired, simultaneous physiological stress to enhance product yield in excess 50% of the total protein or of one gram per liter of the culture. IPTG, lactose and other chemicals known in the art may be used as the agent for chemical induction. These include, inter alia, tryptophan, sugars and analogs, maltose, arabinose, nalidexic acid, oxalinic acid and nitrogen. The physiological stressing agent includes heat shock as well as other

conventional techniques, which may be simulated by the presence of nalidixic acid or other related agents known to induce heat shock genes.

The process of the invention results in a cell lysate where the bulk of the cell lysate is the desired recombinant product, i.e. 10% or less of the recovered protein consists of host protein, and 90% of the protein consists of recombinant protein. The process generally includes a minimum of 2 steps to produce a recombinant protein product from crude cell lysates.

The protein product of this invention is intracellular and not secreted protein (periplasm) as in some other conventional systems. Another unique feature of the invention is that for the bulk of the clones, the protein remains soluble at this level of expression, a major benefit.

The invention demonstrates a new process for growth and induction of bacterial cultures expressing recombinant proteins from the laboratory bench to the pilot scale for production. In previous systems, increased protein turnover and a decrease in the metabolic efficiency of *E. coli* at the high cell densities achieved in the fermentor resulted in failures in scale-up. In these systems, there was a lack of correlation between increased culture volume, increased cell density and increased specific yield of product. These limitations do not occur in the present invention. The data indicate that: 1) yields in shaker flask cultures are directly scalable to the 10 liter fermentor, 2) higher cell densities actually augment the specific yield of ubiquitin fusion proteins, and 3) the

ubiquitin fusion expression clone makes about 3 fold more total protein, including host proteins, than the parent strain of *E. coli*. A series of 3 fermentations with varied temperature shift protocols showed that a maximal heat shock of 12°C (from 30°C to 42°C), done simultaneously with induction, gave a maximal specific yield (over 90% of the total soluble protein by densitometry, 709 mg/liter by protein assay), in which the recoverable ubiquitin fusion product comprised 16% of the wet weight of the cell paste. These results demonstrate the great value of ubiquitin fusion technology for the economical large scale production of peptides.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1: Flow diagram of the ubiquitin fusion process for making recombinant peptides in bacteria.

Fig. 2: Time course and analysis of fermentations D68 and D69. Comparison of growth curves (circles) versus specific yield (diamonds) as measured by densitometric analysis of whole cells samples in SDS-PAGE gels as described in materials and methods. Open markers are D68 datapoints and filled markers are D69 datapoints.

Fig. 3: Time course analysis of protein expression for fermentation D72. Culture samples were taken at one hour time points prior to induction (samples 1-4) and half hour time points following induction (samples 5-9.) Samples were

processed and SDS-PAGE gels were run as described in experimental methods. Each time point is represented by three lanes; "W" for whole cells, "L" for lysate, and "H" for heat supernatant. Ubiquitin standards are shown in lanes marked "U" and protein standards are shown in lanes marked "S". Lane 15 contains an "L" sample from the last timepoint prior to harvest for D81.

Fig. 4: Growth curves for fermentations D72, D73, D74, D80 and D81. Optical density at 600 nm is plotted as a function of time. Cultures were induced at time zero with 1 mM IPTG. The solid line represents a regression through D72, D80 and D81 datapoints.

Fig. 5: Comparison of time courses of specific yield estimates for fermentations D72, D73, and D74 derived by densitometry on whole cell samples. "W" (panel A) versus clarified lysate samples "L" (panel B).

Fig. 6: Comparison of specific yield estimates for fermentations D72, D73, and D74. Derived by densitometry on whole cell samples "W", clarified lysate samples "L", and heat supernatants "H". Open spaces at the tops of bars represent +1 standard deviation from the mean.

Fig. 7: Comparison of soluble protein recovered from time point samples in fermentations D72, D73, D74, D80 and D81. Protein concentration was quantitated in each "L" sample as

described and plotted as a function of the cell density at the time each sample was taken. The protein/O.D. values shown in the inset represent the average and standard deviation of the final five datapoints for each fermentation.

Figure 8: Growth curves for 10 liter fermentations in which ubiquitin peptide fusions were expressed. Fermentation runs of different fusion proteins are designated as shown in Table 4.

Figure 9: SDS-PAGE analysis of the bacterial lysates. Lane 1: ubiquitin standard, 4 μg (Sigma), lanes 2, 4, 6, 8, and 10 contain 3 μl of a 1:10 dilution of each lysate in order as follows: D95, D85, D86, D106, D96, lanes 3, 5, 7, 9, and 11 contain UBP digests of the preceding lysate. Lane 12 contains a mock digest of D96 without the UBP. Lane 13 contains a molecular weight standard (Novex).

Figure 10: HPLC analysis of the D85 bacterial lysate, cleavage reaction and the peptide product. Reverse phase HPLC was done with a gradient of 25-40% buffer A to B run over 18 minutes. Panel A shows an ubiquitin standard (Sigma) elution profile. In panel B, a 1:10 dilution of the D85 bacterial lysate was mixed with an ubiquitin standard, showing the relative purity of the ubiquitin fusion protein, and its location in the gradient elution relative to ubiquitin. Panel C

shows the cleavage reaction of the D85 lysate, with ubiquitin and a clean peptide (see arrow) as products of a complete reaction. Panel D shows the purified peptide product analyzed in 15-30% gradient run over 30 minutes.

Figure 11: FAB Mass Spec Analysis of the peptide product. Panel A shows the "chemical average" of the product while panel B shows a high resolution scan. The product corresponds exactly to the predicted molecular weight of the peptide. FAB mass spec has been used to verify that the molecular weights of the peptides released from the D86 and D95 fusions are correct. The predicted molecular weight of the purified D85 peptide complexed with TFA is 2089.15 Daltons. The peptide peak is shown with an arrow.

Figure 12: A. Plasmid PTre99A. B. Nucleotide sequence encoding human ubiquitin (SEQ ID NO:1).

DETAILED DESCRIPTION OF THE INVENTION

Materials and methods

Bacterial Strains and Plasmids. The host strain, *E. coli*, DH5αF'IQ, was purchased from Life Technologies, Inc., and transformed with the ubiquitin expression vector, pPX2. pPX2 is a derivative of pDSUb that contains a 17 amino acid C-terminal extension of ubiquitin. pDSUb is a derivative of pDS78/RBSII that expresses ubiquitin (25). pPX2 was constructed by inserting a PCR product encoding the peptide and fusion junction in

a directional cloning into pDSUb that had been digested with *Afl*II and *Pst*I, purchased from New England Biolabs and used according to their instructions. Transformants were selected with 100 μg/ml ampicillin, grown for two days at 30°C. The ubiquitin fusion is overexpressed by the addition of IPTG which titrates the lac repressor (DH5αF'IQ contains an upregulated lacI^q gene on the F' factor which is selected with 10 μg/ml neomycin.) Transformants are screened in 10 ml cultures for expression of a fusion protein of the appropriate size. Plasmid DNA is subsequently prepared from clones that exhibit an overexpressed, induced ubiquitin fusion protein, and sequenced by the dideoxy method using the Sequenase Version 2.0 kit (United States Biochemical.) Positive clones are then frozen down and stored. Clone A37, used in the Examples, was maintained on LB plates containing ampicillin (100 μg/ml), neomycin (10 μg/ml) and 1% glucose, at 30°C. It is streaked weekly for up to 10 passages, after which a fresh streak is taken from a frozen seed vial for serial culture, to insure strain authenticity.

A New transformed strain of *E. coli*

This invention includes transformants having unexpectedly better characteristics than parent strains:increased protein content over parent strain, increased expression levels and atypical growth curves.

These unique characteristics are evidenced by the following findings:

1. Transformation frequency in cloning experiments is 10^5-10^6 times lower than expected for the amount of DNA used and the level of competence of the cells. This discrepancy of 10^6 is the same as the frequency with which random single mutations occur in *E. coli*.

2. Since ubiquitin is toxic to the bacteria, any strain that overproduces fusions is somehow tolerized to the intracellular presence of ubiquitin. In fact, the ubiquitin-EGF2 expressing plasmid in A 37 (pPX2) is very stable, even when the ampicillin selection is removed. Attempts to "cure" four of the overexpressing strains of their ubiquitin fusion expression plasmids were made. The greatest overexpresser, the Ub-BMP clone, could not be cured.

3. After curing the EGF2 expression strain we transformed in full length EGF and uteroglobin expression plasmids. These were expressed better in the cured strain than the original clones.

4. In contrast to the parent strain, the expression strain does not grow when a colony on rich media is used to inoculate minimal liquid media with glucose as the only carbon source.

5. The fermentation profiles of the EGF2 strain for expression level and growth are atypical. The parent strain does not exhibit the same "behavior".

6. The EGF2-producing strain exhibits a greater protein content than does the parent strain. This is shown in Fig. 7, which displays soluble

protein as a function of time during the fermentations. For the parent strain the protein content is a flat line, while the protein content increases dramatically for the ubiquitin fusion EGF2 strain. The gel data indicate that this difference is not due simply to the induced expression and production of the ubiquitin fusion, but that host proteins are elevated as well.

Growth conditions. Experiments with A37 were performed in 14 L New Brunswick Microferm bioreactors (10 L w.v.) at the UMD Bioprocess Scale-up Facility. 2X YT medium without glucose, was utilized in all fermentations and was prepared as described by Sambrook (26). In order to provide homogeneous plasmid-bearing cell populations, ampicillin (100 μg/ml) was added to the media. 500 ml starter cultures in 1 L shake flasks were grown from single colonies by overnight incubation (16-20 hours) in LB media (prepared according to Rodriguez and Tait (27)) with ampicillin (100 μg/ml), neomycin (10 μg/ml) and 1% glucose. Shake flask cultures were maintained at 30°C and 200 rpm in a New Brunswick reciprocating environmental shaker. Bioreactors were monitored and controlled to pH 7.0. Dissolved oxygen was maintained above 30% of air saturation. The inoculum used for all experiments was approximately 5% v/v. Cells were induced by IPTG addition to 1 mM at OD_{600} = 2.0. In some experiments, the temperature was raised to 42°C simultaneously with induction.

Analytical Measurements. Cell Mass Density Determination: A Bausch & Lomb Spec 21 DV spectrophotometer was used to measure the optical

density (O.D.) of the samples at 600 nm. Samples above the linear range (0-0.25 OD_{600} units) were diluted with deionized water.

Protein gels. One milliliter samples of growing cultures were taken at half hour intervals throughout each fermentation. The cells were pelleted with a 5 min. centrifugation in a Sorvall MC 12V microfuge at a setting of 6. Cell pellets were then frozen on dry ice and stored at -80°C until processing. To generate protein samples for SDS-PAGE electrophoresis, the cell pellets were rapidly thawed, resuspended in 300 ml of lysis buffer (50 mM Tris, 1 mM EDTA, pH 7.5), and a 100 ml aliquot of the cell suspension was removed to a new tube. The 100 ml aliquots were prepared directly for electrophoresis with the addition of 200 ml of 2X loading buffer (100 mM Tris-HCl pH 6.8, 200 mM DTT, 4% SDS, 0.2% bromophenol blue, 20% glycerol), labelled "whole cells", and heated to 95°C for 15 minutes. The remaining 200 ml of cell suspension was subjected to three cycles of freeze-thaw, alternating between a dry ice-isopropanol bath and a 50°C water bath. Samples were left in the water bath only up to the point at which they became completely thawed, then were immediately frozen again. Following the last freeze-thaw cycle, the samples were placed in a Branson 3200 sonic cleaner bath at maximum setting for 10 min. After sonication, the samples were split into two 100 ml aliquots. One set of these aliquots was placed in the microfuge at maximum setting for 10 min at 4°C. The supernatants of this set were transferred to fresh tubes and the pellets discarded. These

samples were labelled "lysate" and were prepared for electrophoresis with the addition of 200 ml of loading buffer. Five 1 ml aliquots from each sample were removed for protein determinations prior to addition of loading buffer. The last set of aliquots were incubated at 85°C for 10 min prior to centrifugation in the microfuge to remove precipitated host proteins and cellular debris. These were labelled "heat supernatant" and were prepared for electrophoresis with the addition of 200 ml of loading buffer. In this manner, comparable sets of samples from each time point, based on equivalent volumes, were generated.

SDS-PAGE gels were run using a two buffer tricine system in a Novex minigel apparatus, with 10-20% tricine gels purchased from Novex. Anode (bottom) buffer was 0.2 M Tris-HCl, pH 9.0. Cathode (top) buffer was 0.1 M Tris, 0.1 M Tricine, 0.1% SDS, pH 8.25. The protein molecular weight marker, "Multi-Mark", was obtained from Novex. Bovine ubiquitin, used as a standard, was purchased from Sigma. Gels were run at a constant current of 4 mA until the dye marker reached the bottom of the gel. Whole cell samples were loaded in 5 ml volumes, lysates and heat supernatants were loaded in 10 ml volumes. Gels were stained with 0.25% Coomassie Blue R250 (Sigma) and acetic acid:methanol (10%:40%) and destained in the same solution minus the dye.

Densitometry. Stained gels were scanned with a black and white Hewlett Packard Deskscanner 2P with Deskscan 2.0 software on a MacIntosh IIvx computer. Images were analyzed with

shareware "NIH Image" version 1.55. Specific yield data (expressed in percent) represents the area under the ubiquitin fusion peak compared to the total area under all peaks in the same lane.

Protein assays. Protein concentrations were done using the Pierce BCA assay kit, using a 30 min incubation at 37°C, according to the manufacturer's instructions. Protein determinations in all data reflect the concentrations found for designated samples in a fixed sample volume of 5 ml from each of the samples, prior to the addition of SDS-PAGE buffer as described above. Purified bovine ubiquitin (Sigma) was used to generate standard curves in all protein determinations.

Analytical RP HPLC was done on a 4.6 X 250 mm Vydac C18 column with 5 μm particle size. Gradients of buffers A to B (buffer A:0.1% TFA, buffer B:0.1% TFA in acetonitrile) were run either 25-40% over 18 minutes or 15-30% over 30 minutes were done.

Enzymatic reaction. Five hundred mls of the D85 lysate, containing approximately 14 grams of ubiquitin fusion protein, was brought up to 800 mls in UBP reaction buffer (50 mM Tris, 10 mM DTT, 1 mM EDTA, pH 8.5). Approximately 2 mg of UBP enzyme was added and the reaction incubated for 8 hours at room temperature. The reaction was stored at 4°C until preparative RP-HPLC could be done.

Recovery of purified peptide. Preparative RP HPLC was done on a Biotage "kiloprep" instrument. 250 mls of the cleavage reaction was loaded and a 15-40% A:B gradient was run over 45 minutes on a 15 x 7.5 cm column packed with vydac C18 resin (particle size of 15 microns.) Peptide was recovered in 640 mls of solvent and analyzed by analytical IIPLC.

Quantitation of purified peptide. Several one ml aliquots of purified peptide collected from HPLC were lyphilized and the TFA salts were weighed on a Mettler balance (Model #). The average weight of each aliquot was 1.7 mg corresponding to a yield of 1.088 grams from approximately one third of the cleavage reaction (or the digestion of about 4.38 grams of the ubiquitin fusion). Thus, the purified peptide was recovered in nearly quantitative yield.

FAB Mass spec analysis of the lyophilized peptide. The molecular weight of the lyophilized peptide was verified using FAB mass spec analysis by Dr. Mark Rogers of M-Scan, Inc.

Please note that the starting materials are publicly available and, using the taught cloning criterion and procedures, overexpressing transformants, like A37, are can be duplicated in a reproducible manner.

EXAMPLE 1

A. Construction of an ubiquitin expr ssion vector for expression of *E. coli*.

Plasmid PTre99A was purchased from Pharmacia (Fig. 12A) and digested with restriction enzymes, NcoI and PstI, and the linearized vector, a 4.2 kb fragment, was purified from an agarose gel, using Geneclean (Bio101, Inc.). The nucleotide sequence encoding human ubiquitin shown in Fig. 12B, with codon usage optimized for expression in bacteria per Andersson and Kurland (28), was amplified by PCR with oligonucleotides specific for the 5' and 3' end, and containing restriction sites, NcoI and PstI. The PCR fragment was digested and gel purified as described for the vector. The vector and ubiquitin fragments were then ligated and the ligation mix was transformed into DH5αFIQ competent cells obtained from BRL. Several colonies were screened for the insert by restriction digests of plasmid preparations. Plasmids with the correct restriction map were verified by DNA sequencing of the ubiquitin encoding DNA. Correct clones were then screen for this IPTG-inducible expression of ubiquitin. One of the ubiquitin-expressing clones was selected as the master clone and frozen down. The new ubiquitin expression vector was called pPXI.

This parental expression vector is used for construction of new ubiquitin fusions for expression in *E. coli*. The target peptide is cloned as a fusion with ubiquitin into pPXI by synthesizing DNA oligonucleotides, encoding the target peptide, using bacterial codon usage according to reference 28. The other essential features of the oligonucleotide sets are that they contain sequences necessary for the fusion with ubiquitin at the 5' end, a stop condon (TAA) at 3' end, and restriction

sites for cloning. A minimum of 2 oligonucleotides is required for short peptide extensions of ubiquitin. Longer peptides require the synthesis of 4-8 oligonucleotides, up to 75 bases in length, and staggered such that they anneal together at the ends to form a single continuous coding sequence for the target peptide when ligated. Still larger fusions, of proteins (greater than 100 amino acids), with ubiquitin must be constructed stepwise using synthetic oligonucleotides or amplified by RT-PCR from a natural source of the coding sequence. These protein coding sequences are typically amplified with a set of mRNA-specific primers that flank the coding sequence. Following gel purification of the PCR product, a second amplification is done with primers hearing the essential features for cloning of an ubiquitin fusion as described above.

B. Construction of an ubiquitin expression vector for expression in yeast.

Plasmid pPIC9 was purchased from InVitrogen and digested with restriction enzymes. XhoI and EcoRI, and the linearized vector, a kb fragment was purified from an agarose gel, using Geneclean (Bio101, Inc.). The nucleotide sequence encoding human ubiquitin, with codon usage optimized for expression in bacteria per Andersson and Kurland (28), was amplified by PCR with oligonucleotides specific for the 5' and 3' end, and containing restriction sites, XhoI and EcoRI. The PCR fragment was digested and gel purified as described for the vector. The vector and ubiquitin fragments were then ligated and the ligation mix was transformed

into DH5αF'IQ competent cells obtained from BRL. Several colonies were screened for the insert by restriction digests of plasmid preparations. Plasmids with the correct restriction map were verified by DNA sequencing of the ubiquitin encoding DNA. Correct clones were then transformed into the yeast *Pichia pastoris*, and screened for the methanol-inducible expression of ubiquitin. One of the ubiquitin-expressing plasmids was selected as the master clone and frozen down. The new ubiquitin expression vector was called pPX200. This parental expression vector is then used for construction of new ubiquitin fusions for expression in yeast.

C. **Construction of an ubiquitin expression vector for expression in insect cells.**

Plasmid pVL1393 was purchased from InVitrogen and digested with restriction enzymes, EcoRI and PstI, and the linearized vector, a 9.6 kb fragment was purified from an agarose gel, using Geneclean (Bio101, Inc.). The nucleotide sequence encoding human ubiquitin, with codon usage optimized for expression in bacteria per Andersson and Kurland (28), was amplified by PCR with oligonucleotides specific for the 5' and 3' end, and containing restriction sites, EcoRI and PstI.

EXAMPLE 2

Time course of ubiquitin fusion expression during fermentations. Initial fermentation experiments were designed to approximate shaker flask protocols as closely as possible in a 10 liter fermenter. This entailed

inoculation with a fresh 500 mL overnight culture (5% of the expression culture volume), growth at 37°C throughout the fermentation, and induction of ubiquitin fusion expression at O.D.$_{600}$=0.5 [optical density of the bacterial culture read at 600 nm=0.5 (or about 5x10^7 cells/ml)] with 1 mM IPTG. One milliliter culture samples were taken at hourly intervals after induction and specific yield was measured at these time points by densitometric scans of SDS-PAGE gels loaded with whole cells. Two fermentations were run, one with a glycerol feed and one without. Fig. 2 shows the profiles of these fermentations. Both cultures reached a maximum specific yield over 50% of the total cell protein, between 2 and 3 hours post induction, after which cell lysis occurred and specific yield dropped significantly. There was no significant benefit from the glycerol feed, although it was not optimized. All future fermentations with this clone were harvested after two hours of induced expression, prior to lysis.

EXAMPLE 3

Effects of temperature shifts on expression of ubiquitin fusions. During the course of optimizing expression of ubiquitin fusions in shaker flask cultures, it was determined that temperature shifts, either from the inoculum to the expression culture, or during the expression, upon induction, improved overall yields. Much higher yields were obtained when the expression is done at 37°C versus 30°C, even when 1 mM IPTG was used in the 30°C culture and 5 mM IPTG was used in the 37°C

culture. Five fermentations were performed to test the effects of temperature shifts on ubiquitin fusion expression, as summarized in Table 1.

All overnight inoculum cultures were grown in shaker flasks at 30°C. Three fermentation protocols were used. D72 represents the typical expression conditions, in shaker flasks, which were done at a uniform 37°C throughout. D73 was heat shocked by 5°C (from 37°C to 42°C) upon induction with 1 mM IPTG. D74 was heat shocked by 12°C (from 30°C to 42°C) upon induction with 1 mM IPTG. An uninduced control fermentation of A37 (D80) was also run to examine the effect of the presence of the vector on growth, without IPTG induction. A control fermentation, D81, was done with the parent strain DH5αF'IQ, in which a 12°C heat shock upon addition of 1 mM IPTG was done to simulate the most extreme conditions to which the test strain was subjected.

Table 1

Fermentation #	Strain	Initial Temperature	Final Temperature	$O.D._{600}$ at Induction
D72	A37	37°C	37°C	2.0
D73	A37	37°C	42°C	2.0
D74	A37	30°C	42°C	1.9
D80	A37	30°C	37°C	Not Induced
D81	DH5αF'IQ	30°C	42°C	2.0

In Fig. 4, the optical density is depicted over time for all fermentations. From the slope, it is apparent that the growth rate prior to induction

was uniform. In the uninduced fermentation, D80, and in the induced fermentation of the parent strain with no plasmid, D81, the final O.D. was approximately 5.4 units. However, the growth rate was reduced during the induction phase of the induced cultures overexpressing ubiquitin. Moreover, as the degree of induction or stress was increased, the growth rate decreased. The most severely affected fermentation was grown at 30°C, induced with IPTG and simultaneously heat shocked to 42°C.

Fig. 3 shows a time course of protein samples taken from D72. For each time point, the gel lanes represent; 1) whole cells ("W"), 2) the supernatant of cell lysates ("L"), and 3) the supernatant of an 85°C incubation of the cell lysate ("H"). Therefore, the "L" and "H" lanes represent recoverable, soluble and heat stable protein. Samples for protein analysis were taken at 1 hour intervals prior to induction and at half hour intervals following induction. Each sample for each time point in each fermentation was processed identically. Protein concentration determinations were done on recovered protein samples and heat supernatant samples, and these data appeared consistent with the degree of staining in these samples on the gels. Five microliter volumes of whole cell extracts were loaded, while 10 microliter volumes for the "L" and "H" lanes were loaded. Specific yields, calculated as the percent ubiquitin fusion protein in total cellular protein, for each set of samples, were determined by densitometry on whole cell "W" lanes. Lane 15 contains a sample of

the D81 fermentation taken at the end of the fermentation prior to harvest, when the cell density reached a maximum. Using this, a comparison of the host proteins in the parent strain can be made with the A37 expression strain. The lack of protein bands in the parent strain compared to the visible non-ubiquitin protein bands in A37, in identically processed samples, demonstrates a reproducible difference between the strains. Protein assays on these samples indicate a 2 fold difference in the total protein concentration, measured with the BCA assay. Comparable bands for host proteins, such as GroEl, can be visualized in DH5αF'IQ when an 100-fold less dilute sample is loaded per lane than A37 samples.

In Fig. 5, the specific yield based on whole cell protein and specific yield based on soluble protein are plotted. The specific yield based on whole cell protein was relatively insensitive to the induction conditions. However, the yield of soluble proteins increased when heat shock was applied. The lowest yield (~65%) was found in the fermentation induced by IPTG addition at 37°C, D73. The culture grown and induced at 37°C but also heat shocked to 42°C, D73, resulted in slightly higher yield (~72%). In the fermentation grown at 30°C and simultaneously induced and heat shocked, D74, the soluble fraction of ubiquitin was effectively 100%. There was no significant residual protein in any PAGE gels of samples from this experiment.

Table 2 illustrates the relative advantages of the heat shock protocols. There is a

clear benefit to heat shock, simultaneous with induction of the ubiquitin fusion protein. Although the overall yield of biomass is lowest for the maximally shocked culture, D74, it yielded the highest amount of ubiquitin fusion protein, at 709 mg/liter, constituting 16% of the wet weight of the cell paste.

TABLE 2						
Fermentation	A	B	C	D	E	F
D72	12.6	0.76	52	.390	55	7.1
D73	11.1	0.66	74	.491	39	12.6
D74	12.4	0.75	95	.709	44	16.1
D80	6.9	0.42	82	.340	71	4.8

A - Micrograms of protein measured by BCA assay in 5 microliter of each "L" sample for the last sample in each run.

B - Conversion of column A to micrograms of protein per microliter of bacterial culture, taking dilution factors into account.

C - Percent ubiquitin fusion, versus total protein, estimated by densitometry in the corresponding "L" sample.

D - Specific yields of ubiquitin fusion in grams per liter of bacterial culture (B x C x 1000 mg/g).

E - Wet weight of the cell paste harvested from the 10 liter cultures in grams.

F - Percent of wet weight of cell paste comprised of the ubiquitin fusion.

EXAMPLE 4

Effect of *in vivo* heat shock versus *in vitro* heat shock on protein recovery. Fig. 6 shows the effect of heat shock on the specific yield of the ubiquitin fusion protein, relative to total protein, in whole cells, lysed cell supernatants

(soluble protein) and heat supernatants (heat stable protein). As previously noted, the specific yields in whole cells and heat supernatants were roughly equivalent for all three growth and induction protocols. However, the specific yields of ubiquitin fusion protein in the soluble protein extracts increased with the severity of the heat shock upon IPTG induction. D74 exhibited essentially 100% ubiquitin fusion protein in the soluble extracts, about 30% more than the other two fermentations. Curiously, the specific yield of the ubiquitin fusion decreased by about 20% in the D74 sample following the *in vitro* heat shock. The practical implication of this observation is that the use of an *in vivo* heat shock is more effective in generating higher overall yields of soluble ubiquitin fusion protein than the *in vitro* purification step involving an 85°C incubation followed by centrifugation. Thus, the use of an *in vitro* heat precipitation step is unnecessary when an *in vivo* heat shock is done.

EXAMPLE 5

Comparison of parent and ubiquitin fusion *E coli* strains. A significant difference in protein

metabolism between the strain expressing the ubiquitin fusion, A37, and its parent strain, DH5αF'IQ, is apparent from our data, as shown in Fig. 7. A comparison of protein content from identically processed samples for A37 and its parent strain show that as the cell density of the culture increases in A37, the protein per unit volume also increases, while the protein per unit culture volume decreases as the cell density increases for the parent strain. Therefore, the protein content per cell decreases with increasing cell density for the parent strain, reflecting the loss in metabolic efficiency typically seen in high density cultures (9). However, uninduced A37, in fermentation D80, shows a similar increase in protein with cell density. Induced A37, in fermentations D72-74, shows an increase in protein greater than the increase in cell density. The uninduced A37 (D80) contains three times (0.5 mg/ml vs. 0.17 mg/ml) as much total soluble protein as the parent strain (D81) at harvest, and the induced A37 strain (in D72, D73, and D74) yielded between 4.3-5.0 times as much soluble protein as the parent strain. The total amount of soluble protein is dependent on the

presence of the ubiquitin fusion protein as well as on the induction of the ubiquitin fusion protein.

EXAMPLE 6

Fermentation scale up. Model ubiquitin peptide fusion clone (A37) was grown in a 10 liter fermentor to high cell density at 37°C prior to IPTG induction of the ubiquitin fusion protein, in a rich "superbroth" (26). The conditions used were identical to those for the initial fermentation experiments with this strain. Fig. 8 shows a typical growth curve for this strain. The cells were harvested by centrifugation 2.5 hours after induction with 1 mM IPTG. This generated 175 grams of cell paste, which was processed as a single batch for recovery of the ubiquitin fusion protein.

Recovery of the ubiquitin fusion protein. Soluble protein was recovered from the cell paste after: 1) resuspending in 600 ml of lysis buffer (50 mM Tris, pH 10), 2) 3 freeze-thaw cycles, alternating between a dry ice ethanol bath and a 50°C water bath, 3) followed by sonication in a Branson Model 3200 sonic cleaner bath at maximum setting for 15 min, and 4) centrifugation at 10K for 20′ in a Sorvall RC-5B GSA rotor. The supernatant

("first extract", approx. 700 ml) was transferred to a 1 liter plastic bottle for storage at -80°C. The pellet was resuspended in another 300 ml of lysis buffer and the extraction process was repeated. The supernatant from the second procedure was labelled "second extract". The second pellet was also frozen for possible future extractions.

Characterization of the soluble protein extracts. BCA assays were also done to quantitate the soluble protein, using a ubiquitin standard (Sigma) to generate the standard curve. A total of 14 grams of protein was recovered in the first extract and 3 grams of protein was recovered in the second extract. Spectral analysis indicated that there was a significant amount of DNA contamination, approximately 110 mg/ml in each sample. Fig. 9 shows an SDS-PAGE gel run on the first and second extracts. Both extracts contain virtually pure ubiquitin fusion protein with very little contaminating host proteins. This result is consistent with protein recoveries in previous fermentations in which the ubiquitin fusion is by far the predominant soluble protein at this stage of the fermentation. Nearly 17 grams of ubiquitin fusion protein were recovered from 175 grams of wet

cell paste, representing almost 10% of the wet cell weight.

Fig. 10 shows HPLC traces for the D0085 "first extract" plus a ubiquitin standard (panel A) and a test reaction of D0085 "first extract" with the cleavage enzyme, UCH-L3 (panel B). Not only is the purity of the simple cell lysate good, in terms of contaminating proteins, but the crude lysate is a good substrate for UCH-L3. The reaction goes to completion and the peptide is not degraded.

EXAMPLE 7

Bulk UCH-L3 reaction. 500 ml of the D0085 "first extract" was brought up to 800 ml with UCH-L3 reaction buffer (50 mM Tris, 10 mM DTT, 1 mM EDTA, pH 8.5). Approximately 2 mg of UCH-L3 (1 ml of bacterial lysate containing UCH-L3 in catalytic amounts) was added and the reaction incubated for 8 hours at room temperature. The reaction was checked for completion by SDS-PAGE, and this showed a small quantity of uncleaved fusion remaining. Another 3 mg of UCH-L3 was added and the reaction was incubated at 30°C for another 4 hours. The reaction was stored at 4°C between incubations where the

enzyme continues to cleave peptide product from the ubiquitin fusion partner.

Scale purification of peptide. One third of the cleavage reaction (about 250 ml) was loaded onto a 15 x 7.5 cm column packed with Vydac C18 resin (particle size of 15 microns) and run on a Biotage "kiloprep" instrument. A TFA:acetonitrile gradient was then run. Peptide was recovered in 1 liter of solvent and analyzed by analytical HPLC. Fig. 11 shows the chromatogram of the peptide. The degree of purity is excellent and nearly 100% of the peptide product is recovered.

EXAMPLE 8

Four different ubiquitin peptide fusion clones, ranging in size from 8 to 70 amino acids, were grown in 10 liter fermentation batches.

Growth conditions. "Super terrific" broth (Sambrook, 1989; UMD modification) (ref 26) was utilized in all fermentations. 500 ml starter cultures in 1 L shake flasks were grown from single colonies by overnight incubation (16-20 hours) in LB media (prepared according to Rodriguez and Tait (1983)) with ampicillin (100 μg/ml), neomycin (10 μg/ml) and 1% glucose. Shake flask cultures were

maintained at 30°C and 200 rpm in a New Brunswick reciprocating environmental shaker. Bioreactors were monitored and controlled to pH 7.0. Dissolved oxygen was maintained above 30% of air saturation. The inoculum used for all experiments was approximately 5% v/v. Cells were induced by IPTG addition to 1 mM at OD_{600}-5.0.

Enzymatic reaction. Five hundred ml of the D85 lysate, containing approximately 14 grams of ubiquitin fusion protein, was brought up to 80 ml in UBP reaction buffer (50 mM Tris, 10 mM DTT, 1 mM EDTA, pH 8.5). Approximately 2 mg of UBP enzyme (UCH-L3) was added and the reaction incubated for 8 hours at room temperature. The reaction was stored at 4°C until preparative RP-HPLC could be done.

Recovery of purified peptide. Preparative RP HPLC was done on a Biotage "kiloprep" instrument. 250 ml of the cleavage reaction was loaded and a 15-40% A:B gradient was run over 45 minutes on a 15 x 7.5 cm column packed with Vydac C18 resin (particle size of 15 microns). Peptide was recovered in 640 ml of solvent and analyzed by analytical HPLC.

Quantitation of purified peptide. Several one ml aliquots of purified peptide collected from HPLC were lyophilized and the TFA salts were weighed

on a Mettler balance. The average weight of each aliquot was 1.7 mg, corresponding to a yield of 1.088 grams from approximately one third of the cleavage reaction (or the digestion of about 4.38 grams of the ubiquitin fusion). Thus, the purified peptide was recovered in nearly quantitative yield.

FAB Mass spec analysis of the lyophilized peptide. The molecular weight of the lyophilized peptide was verified using FAB mass spec analysis. Growth curves for these fermentations are shown in figure 8. The recombinant fusion proteins were expressed and cell pastes harvested. Induction of fusion protein synthesis was done at relatively low cell densities (O.D.600=5.0).

Wet cell paste was processed as described and bacterial lysates generated by repeated freeze-thaw cycles. These were analyzed by SDS-PAGE and the lysates were found to be very concentrated. Figure 9 shows an SDS-PAGE gel where each sample lane contains 3 μl of lysate diluted 1:10 with water. These lysates contain at least 95% ubiquitin fusion protein by densitometry. The D106 yield is low due to toxicity of the ubiquitin fusion to host strain, resulting in low biomass recovered. The major contaminant appears to be DNA, measured at

100-115 μg/ml. Estimated yields of ubiquitin fusion protein for each fermentation run are shown in Table 3.

TABLE 3
YIELDS OF FUSION PROTEINS

Run #	Length of Peptide	Grams of cell paste	Grams of fusion protein
D85	17	175	17-52
D86	33	135	29-58
D95	8	197	60-120
D96	70	247	43-86
D106	53	35	1.5

Conclusions

1. Ubiquitin-peptide fusions, ranging in size from 8 to 70 amino acids are expressed at very high levels in *E. coli*. In this system, more than 10% of the wet weight of the cell paste is fusion protein. These fusion proteins are soluble and are recovered easily in simple freeze-thaw lysates. This is the highest reported recovery of a recombinant protein from *E. coli*.

2. The crude extract generated by fermentation of ubiquitin-peptide fusions is extremely clean, and requires little, if any, yield-cutting downstream processing, prior to the enzymatic release of the peptide extension.

3. A simple three-step process has been demonstrated with near quantitative yields of purified peptide product from bacterial cell pastes:

1) freeze-thaw lysis, 2) enzymatic release of peptide from ubiquitin, and 3) purification of the peptide by RP-HPLC.

The sequences of the peptides of Table 3 are shown in Table 4.

TABLE 4

Size of Extension	Level of Overexpression	Amino acid sequence
8 amino acids	~80% of total protein	KGDEESLA (SEQ ID NO:2)
17 amino acids	~78% of total protein	NCVVGYIGERCQYRDLK (SEQ ID NO:3)
17 amino acids	~73% of total protein	NCVVGYIGERCQYRDLA (SEQ ID NO:4)
22 amino acids	~70% of total protein	NCVVGYIGERCQYRDLKWWELR (SEQ ID NO:5)
22 amino acids	~88% of total protein	NCVVGYIGERCQYRDLKWWELA (SEQ ID NO:6)
23 amino acids	~38% of total protein	GIGKFLHSAKKFGKAFVGIEMNS (SEQ ID NO:7)
28 amino acids	~70% of total protein	SLRRSSCFGGRMDRIGAQSGLGCNSFRY (SEQ ID NO:8)
29 amino acids	~72% of total protein	HSQGTFTSDYSKYLDSRRAQDFVQWLMNT (SEQ ID NO:9)
33 amino acids	~75% of total protein	CGNLSTCMLGTYTQDFNKFHTFPQTAIGVGAPA (SEQ ID NO:10)
53 amino acids	~50% of total protein	NSDSECPLSHDGYCLHDGVCMYIEALDLYACNCVVGYIGERCQYRDLKWWELR (SEQ ID NO:11)
70 amino acids	~70% of total protein	EICPSFQRVIETLLMDTPSSYEAAMELFSPDQDMREAGAQLKKLVDTLPQKPRESIIKLMEKIAQSSLCN (SEQ ID NO:12)

The lysates also were evaluated by reverse phase HPLC. All lysates are efficiently digested to ubiquitin plus product. Figure 10 shows the results obtained for the D85 lysate. The enzymatic cleavage reaction is measured by RP-HPLC, shown in panel C.

Analytical reactions typically go to completion at room temperature in 20 minutes. The bulk cleavage of D85 was done over a period of several hours at a weight ratio of 1:7000, enzyme:substrate. Crude lysates appeared to be good enzymatic substrates, as evidenced by cleavage reactions in Fig. 9 (lanes 3, 5, 7, 9, 11). A major advantage is that the reaction goes to completion and the peptide is not degraded. The reaction in lane 11 goes to completion with an extended incubation time of 120 minutes. This reaction is slower due to the large size of the peptide extension of this fusion (70 amino acids).

The peptide product of D85 was then purified by preparative RP-HPLC. Figure 10 shows the RP-HPLC analysis of the purified peptide. It is >95% pure and was recovered in high yield, corresponding to 302 mg/liter of bacterial culture.

FAB mass spec was used to determine the molecular weight of the purified peptide (Figure 11). The product corresponds to the predicted molecular weight of the TFA salt of the peptide. FAB mass spec has been used to verify that the molecular weights of the D86 and D95 peptides are correct, as well.

Analytical work shows that the correct product is obtained with this process. A preparative purification of one product demonstrates that the method produces an exceptionally high product yield. This work demonstrates the value of ubiquitin fusions as a peptide production method with potential for obtaining much higher quantities

of peptide from 10 liter fermentations than has heretofor been possible.

DISCUSSION

Ubiquitin fusion technology is a powerful bioprocess for producing proteins and peptides in the bacterium *Escherichia coli*. The two components of the technology are 1) a ubiquitin-specific protease, which cleaves only C-terminal extensions from ubiquitin, and 2) an expression system consisting of a suitable *E. coli* host strain paired with a plasmid that encodes the ubiquitin fusion. Components have been put together to produce several peptides, ranging in size from 8 to 70 amino acids. A combination of the exceptional over-expression of the ubiquitin-peptide fusion protein in the bacteria and the simple, three step purification, including the enzymatic cleavage reaction, insures cost-effective production of gram quantities of peptides. This technology has the capacity to generate over 2 grams of pure peptide from 12 grams of ubiquitin fusion, using the cell paste generated from a single 10-liter fermentation.

The value of a bioprocess lies in its efficiency. In order to evaluate overall efficiency and make general comparisons to other bioprocesses, two numbers that effect overall yield are required. These are: 1) the yield of raw material that can be generated by manipulation of growth of the organism and expression of the protein product ("upstream processing"), and 2) the percent losses incurred in the purification of the final product ("downstream processing").

Multigram yields for upstream processing using ubiquitin fusion technology have been shown for a significant size range of C-terminal extensions of ubiquitin. -- The outstanding yields for these fusions illustrate that this process is quite generic and could be used for the production of peptides up to at least 70 amino acids. These are the highest reported recoveries of recombinant proteins from *E. coli*.

The UCH-L3 ubiquitin-specific protease is capable of cleaving very large amounts of fusion protein at very high substrate to enzyme ratios (7000:1), under ordinary benchtop conditions. See Recksteiner, M. (ed.), Ubiquitin, Plenum Press, N.Y. (1988).

The crude extract generated by fermentation of the model ubiquitin fusion is extremely clean, and requires little, if any, yield-cutting downstream processing. The percent losses of sample in downstream processing are relatively small and nearly 100% of product was recovered as lyophilized peptide from the Biotage column.

There are several unique features of ubiquitin fusion expression in fermentor cultures. Fermentations with ubiquitin expressing bacteria show an unusual growth profile. The first set of fermentations (D68 and D69) showed that once the level of ubiquitin fusion reaches an overwhelming majority of the total bacterial cellular protein, the cells die, and specific yield decreases dramatically in less than 30 minutes. The cells may be lysing from protein overload, or from a lack of sufficient host protein content to mediate

successful cell division. Fermentation parameters, which are monitored routinely, such as cell density, dissolved oxygen, and pH do not predict these losses. Therefore timing of the harvest is critical, and must be done no more than 2.5 hours after induction with 1 mM IPTG.

The D68 and D69 data also show that, contrary to the common experience of losing specific yield productivity in fermentor cultures, this system was equal or superior to the shaker flask results. Densitometry of Coomassie stained SDS-PAGE gels indicated a strong trend towards a vast majority of ubiquitin fusion over host protein content. The results show that the ubiquitin fusion was both soluble and recoverable, and the quantities recovered are consistent with initial estimates obtained by densitometry of SDS-PAGE on whole cells.

Another unique feature of ubiquitin fusion expression is the fact that a heat shock stress actually augments productivity in this system. This phenomenon was initially observed in shaker flask cultures which were grown at 30°C prior to induction in order to maintain vector stability, then shifted to 37°C upon induction for maximal expression. Although the yield in biomass is significantly reduced by the severe heat shock (D74, from 30°C to 42°C), the increase in specific yield of the ubiquitin fusion more than compensates for the loss, for an overall yield of 709 mg/liter, comprising 16% of the wet weight of the cell paste. It is surprising and unexpected that this amount of a recombinant protein can be produced intracellularly and remain soluble, especially when temperature

elevations during expression exacerbates the tendency to form inclusion bodies for many proteins (20A). However, ubiquitin is an eukaryotic heat shock protein (1) and remains soluble for an 85°C heat treatment during purification, as previously reported (7).

The fermentation protocol also effects the downstream purification of the ubiquitin fusion. The heat shock *in vivo* decreased the effectiveness of the *in vitro* heat treatment, with respect to specific yield of the ubiquitin fusion in unheated versus heated lysates. The heat treatment did not appear to decrease the amount of the ubiquitin fusion in these samples, rather the *in vivo* heat shock appears to have increased the yield of host proteins in the heat treated samples versus the unheated samples. It appears that the physiological basis for this qualitative difference in host proteins in cells subjected to heat shock is that the *in vivo* heat shock has altered the host proteins, rendering them more soluble in a warmer environment, and thereby ensuring survival of the bacteria. Finally, the results demonstrate that the protein content of the induced recombinant cultures is significantly higher than that of the parent strain. In the induced A37 cultures, the protein content stays roughly constant as the cell density rises, while in the parent strain, the protein percentage of each cell decreases as the cell density increases. This decrease in protein content may be due in part to the severe stresses of heat shock and/or the chemical induction with IPTG. The decrease in protein content supports the view that

the metabolic efficiency of the cells decreases with increasing cell density (12, 23, 24), and helps to explain the lower specific yields obtained for high density fermentor cultures versus those obtained with shaker flask cultures where densities are much lower.

In contrast, the protein content of ubiquitin fusion strain of the invention is not decreased by the induction or the heat shock. The increase in protein content is due mostly to the accumulation of the ubiquitin fusion, although the host proteins also appear to be present in higher amounts in this strain, relative to the parent strain, based on Coomassie visualization in the gels.

30 peptides ranging in size from 8 to 70 amino acids, have been produced as ubiquitin fusions. Peptides of practical interest (mostly therapeutic applications) shown in Table 4. Theoretically, there is no size limit to the expression system. Compounds that could be produced using the invention include, *inter alia*, peptide hormones, antimicrobial peptides, peptide epitope fragments, ACTH, VIP, GHRH, CGRP, amylin, osteocalcin, insulin, natriuretic peptides.

References

1. Rechsteiner, M. (ed.) Ubiquitin, Plenum Press, New York (1988).

2. Butt, T. R., S. Jonnalagadda, B.P. Monia, E.J. Sternberg, J.A. Marsh, J.M. Stadel, D.J. Ecker, and S.T. Crooke. Ubiquitin fusion augments the yield of cloned gene products in *Esherichia coli*. PNAS 86:2540-2544 (1989).

3. Cherney, B.W., B. Chaudhry, K. Bhatia, T.R. Butt, and M. Smulson. Expression and mutagenesis of human poly(ADP-ribose) polymerase as a ubiquitin fusion protein from *Escherichia coli*. Biochemistry 30:10420-10427 (1991).

4. Baker, R.H., S.A. Smith, R. Marano, J. McKee, and P.G. Board. Protein expression using cotranslational fusion and cleavage of ubiquitin. JBC 269: 25381-25386 (1994).

5. McDonnell, D.P., J.W. Pike, D.J. Drutz, T.R. Butt, and B.W. O'Malley. Reconstitution of the vitamin-D-responsive osteocalcin transcription unit in *Saccharomyces cerevisiae*. MCB 9:3517-3523 (1989).

6. Poletti, A., N.L. Weigel, D.P. McDonnell, W.T. Schrader, B.W. O'Malley, and O.M. Connelly. A novel, highly regulated, rapidly inducible system for the expression of chicken progesterone receptor, cPR_A, in *Saccharomyces cerevisiae*. Gene 114:51-58 (1992).

7. Yoo, Y., K. Rote, and M. Rechsteiner. Synthesis of peptides as cloned ubiquitin extensions. JBC 264:17078-17083 (1989).

8. Enfors, S.O. Control of in vivo proteolysis in the production of recombinant proteins, TIBTECH 10: 310-315 (1992).

9. Georgiou, G. Optimizing the production of recombinant proteins in microorganisms, AICHE J. 34:1233-1245 (1988).

10. Ramirez, D.M., and W.E. Bentley. Enhancement of recombinant protein synthesis and stability via coordinated amino acid addition. Biotech. and Bioeng. 41:557-565 (1993).

11. Ramirez, D.M., and W.E. Bentley. A fed-batch policy to improve foreign protein synthesis and stability by avoiding stress responses. Biotech. and Bioeng. (*in press*, 1995).

12. Majewski, R.A. and Domach, M.M. Simple constrained-optimization view of acetate overflow in *E. coli*. Biotechnol. Bioeng. 35:732-738 (1990).

13. Rinas, U., Kracke-Helm, H.A., and Schugerl, K. Glucose as a substrate in recombinant strain fermentation technology. Appl. Microbiol. Biotechnol. 31:163-167 (1989).

14. Gottesman, S. and Maurizi, M.R. Regulation by proteolysis: Energy-dependent proteases and their targets. Microbiol. Rev. 56:592-621 (1992).

15. Miller, C. Protein degradation and proteolytic modification, pp. 680-691. In: F.C. Neidhardt, J.L., Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger (eds.). *Escherichia coli* and *Salmonella typhimurium*. Vol.1., American Society for Microbiology, Washington, DC. (1987).

16. Neidhardt, F.C., and VanBogelen, R.A., Heat shock response, pp.1334-1345. In: F.C. Neidhardt, J.L., Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger (eds.), *Escherichia coli* and *Salmonella typhimurium*, Vol.1., American Society for Microbiology, Washington, DC. (1987).

17. Harcum, S.W. and Bentley, W.E. Detection, quantification, and characterization of proteases in recombinant *Escherichia coli*, Biotech. Techniques, 7:441-447.

18. Harcum, S.W. and Bentley, W.E., Response dynamics of 26, 34, 39, 54, and 80 kDa proteases in induced cultures of recombinant *Escherichia coli*, Biotechnol.Bioeng.,.--

19. Goff, S.A. and Goldberg, A.L., Production of abnormal proteins in *E.coli* stimulates transcription of lon and other heat shock genes. Cell 41: 587-595 (1985).

20. Labean, T.H., S.A. Kaufman, and T.R. Butt. Design, expression and characterization of random

sequence polypeptides as fusions with ubiquitin. FASEB J. 6:A471 (1992).

21. Parsell, D.A., and S. Lindquist. The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. Annu. Rev. Genet. 27:437-496 (1993).

22. Chen, Y., D. Pioli, and P.W. Piper. Overexpression of the gene for polyubiquitin in yeast confers increased secretion of a human leukocyte protease inhibitor. Bio/Technol. 12:819-823 (1994).

23. Lee, Y. and H.W. Blanch. Defined media optimization for growth of recombinant *Escherichia coli* X90. Biotech & Bioeng. 41:221-230 (1993).

24. Lee, Y. and H.W. Blanch. Recombinant trypsin production in high cell density fed batch cultures of *Echerichia coli*. Biotech. & Bioeng. 41:781-790 (1993).

25. Hochuli, E., W. Bannwarth, H. Dobeli, R. Gentz, and D. Stuber. Biotechnology 6:1321-1325 (1988).

26. Sambrook, J., E.F. Fritsch, and T. Maniatus. Molecular Cloning: A Laboratory Manual (second edition). Cold Spring Harbor Laboratory Press. (1989)

27. Rodriguez, R. L. and Tait, R.C., Recombinant DNA techniques: An Introduction, The Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA (1983).

28. G.G.E. Andersson and C.G. Kurland. "Codon Preferences in Free-Living Microorganisms." Microbiological Review 54:198-210 (1990).

30. D.J. Ecker, et al. "Gene Synthesis, Expression, Structures, and Functional Activities of Site-specific Mutants of Ubiquitin". J. Biological Chemistry 262:14213-14221 (1987).

WHAT IS CLAIMED IS:

1. A host cell comprising an expression vector encoding a ubiquitin-peptide fusion protein, said host cell being capable of expressing said protein in an amount of at least approximately 1 gram per liter of culture medium.

2. The host cell of claim 1 which is a bacteria.

3. The host cell of claim 2 which is E. coli.

4. The host cell of claim 3 which is A37.

5. The host cell of claim 1 which is a eukaryotic cell.

6. The host cell of claim 5 which is a yeast.

7. The host cell of claim 5 which is a insect cell.

8. A transformed strain of E. coli that is capable of expressing a recombinant protein in an amount of at least approximately 1 gram per liter of culture medium.

9. The transformed strain of claim 8 in which said recombinant protein is an ubiquitin fusion protein.

10. A fermentation method for producing a target peptide from an ubiquitin fusion protein comprising:

a) inducing the synthesis of the fusion protein by a transformant contained in a fermentation culture;

b) culturing the transformant under conditions suitable for intracellular production and accumulation of the induced fusion protein;

c) recovering said fusion protein in at least approximately 50% specific yield;

d) hydrolysing said induced fusion protein using an ubiquitin specific hydrolase to release the target peptide, and

e) recovering said target peptide.

11. The method for producing a target peptide from an ubiquitin fusion protein according to claim 10 wherein the culturing conditions selected result in the intracellular production and accumulation of the induced fusion protein in at least 50% specific yield.

12. The method for producing a target peptide from an ubiquitin fusion protein according to claim 10 wherein the culturing conditions selected result in the intracellular production and accumulation of the induced fusion protein in at least approximately one gram of said fusion protein per liter of said culture.

13. The method of claim 10 wherein the transformant is selected from prokaryotic or eukaryotic cells.

14. The method of claim 13 wherein the prokaryotic cell is a bacteria.

15. The method of claim 14 wherein the bacteria is E. coli.

16. The method of claim 15 wherein the E. coli is A37.

17. The method of claim 13 wherein the eukaryotic cell is plant, mammalian, yeast or insect cell.

18. The method of claim 10 wherein synthesis is induced by means of a chemical agent.

19. The method of claim 18 wherein said chemical agent is IPTG, lactose, tryptophan, sugars and analogs, maltose, arabinose, nalidexic acid, oxalinic acid or nitrogen.

20. The method of claim 18 wherein said chemical agent is IPTG or lactose.

21. The method of claim 18 wherein said chemical agent is appropriate for any one of the following promoters T7, lac, lacUV5, tac, trc, trp, lexA, malE, araC, λpPL, and nifA.

22. The method of claim 10, further comprising purifying said fusion protein using a heat step.

23. The method of claim 22, wherein said heat step is carried out at about 85°C.

24. The method of claim 10, further comprising application of a physiological stress agent at the time of inducing synthesis.

25. The method of claim 24, wherein the physiological stress agent is heat shock.

26. The method of claim 25, wherein said heat shock is a temperature increase of about 12° C.

27. The method of claim 18, further comprising application of a physiological stress agent at the time of inducing the synthesis of the fusion protein.

28. The method of claim 10, wherein said target peptide has between 8 and 70 amino acids.

29. The method of claim 10, wherein said target peptide is selected from the group peptide hormones, antimicrobial peptides, peptide epitope fragments, ACTH, VIP, GHRH, CGRP, amylin, osteocalcin, insulin, and natriuretic peptides.

30. The method of claim 10, wherein said target peptide is selected from peptides those having between 2 amino acids and 6000 amino acids.

31. The method of claim 10, wherein said target peptide is selected from peptides those having between 6 amino acids and 500 amino acids.

32. The method of claim 10, wherein said target peptide has between 8 and 70 amino acids.

33. The method of claim 10, wherein said target peptide is selected from peptides those having KGDEESLA, NCVVGYIGEROQYRDLK, NCVVGIYGERCQYRDLA, NCVVGYIGERCQYRDLKWWELR, NCVVGYIGERCQYRDLKWWELA, GIGKFLHSAKKFGKAFVGEMNS, SLRRSSCFGGRMDRIGAQSGLGCNSFRY, HSQGTFTSDYSKYLDSRRAQDFVQWLMNT, CHNLSTCMLGTYTQDFNKFHTFPQTAIGVGAPA, NSDSECPLSHDGYCLHDGVCMYIEALDLYACNCVVGYIGERCQYRDLKWWELR, and EICPSFQRVIETLLMDTFSSYEAASVIELFSPDQDMREAGAQLKKLVDTLPQKPRESIIKLMEIKIAQSSLCN.

34. A kit for the scale up of a fermentation process for producing recombinant fusion ubiquitin-target peptide fusion proteins comprising an expression vector, a nucleotide sequence encoding an ubiquitin and a host cell capable of expressing said fusion protein in an amount of at least approximately 1 gram per liter of culture medium.

35. The kit of claim 33 wherein the host cell is a bacteria.

36. The kit of claim 34 wherein the bacteria is E. coli.

37. The kit of claim 35 wherein the E. coli is A37.

38. The kit of claim 33 wherein the host cell is eukaryotic cell.

39. The kit of claim 37 wherein the eukaryotic cell is a plant, mammalian, yeast or insect cell.

40. The method of claim 10, wherein said target peptide is selected from peptides those having KGDEESLA (SEQ ID NO:2), NCVVGYIGERCQYRDLK (SEQ ID NO:3), NCVVGIYGERCQYRDLA (SEQ ID NO:4), NCVVGYIGERCQYRDLKWWELR (SEQ ID NO:5), NCVVGYIGERCQYRDLKWWELA (SEQ ID NO:6), GIGKFLHSAKKFGKAFVGIEMNS (SEQ ID NO:7), SLRRSSCFGGRMDRIGAQSGLGCNSFRY (SEQ ID NO:8), HSQGTFTSDYSKYLDSRRAQDFVQWLMNT (SEQ ID NO:9), CGNLSTCMLGTYTQDFNKFHTFPQTAIGVGAPA (SEQ ID NO:10), NSDSECPLSHDGYCLHDGVCMYIEALDLYACNCVVGYIGERCQYRDLKWWELR (SEQ ID NO:11), and EICPSFQRVIETLLMDTPSSYEAAMELFSPDQDMREAGAQLKKLVDTLPQKPRESIIKLMEKIAQSSLCN (SEQ ID NO:12).

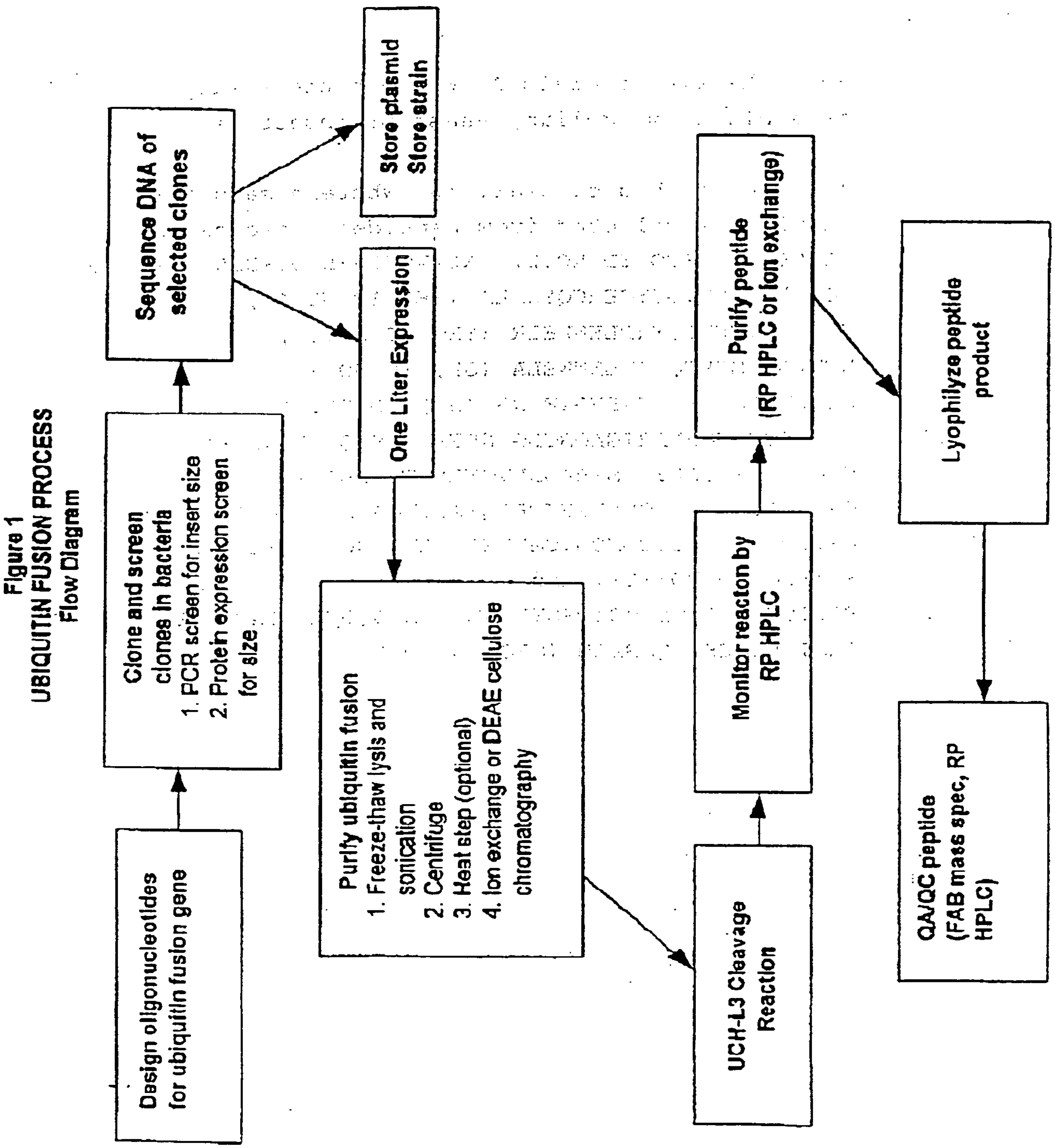


Figure 1
UBIQUITIN FUSION PROCESS
Flow Diagram
Design oligonucleotides for ubiquitin fusion gene
Clone and screen clones in bacteria
1. PCR screen for insert size
2. Protein expression screen for size.
Sequence DNA of selected clones
Store plasmid
Store strain
One Liter Expression
Purify ubiquitin fusion
1. Freeze-thaw lysis and sonication
2. Centrifuge
3. Heat step (optional)
4. Ion exchange or DEAE cellulose chromatography
UCH-L3 Cleavage Reaction
Monitor reacton by RP HPLC
Purify peptide (RP HPLC or Ion exchange)
Lyophilyze peptide product
QA/QC peptide (FAB mass spec, RP HPLC)

FIGURE 1

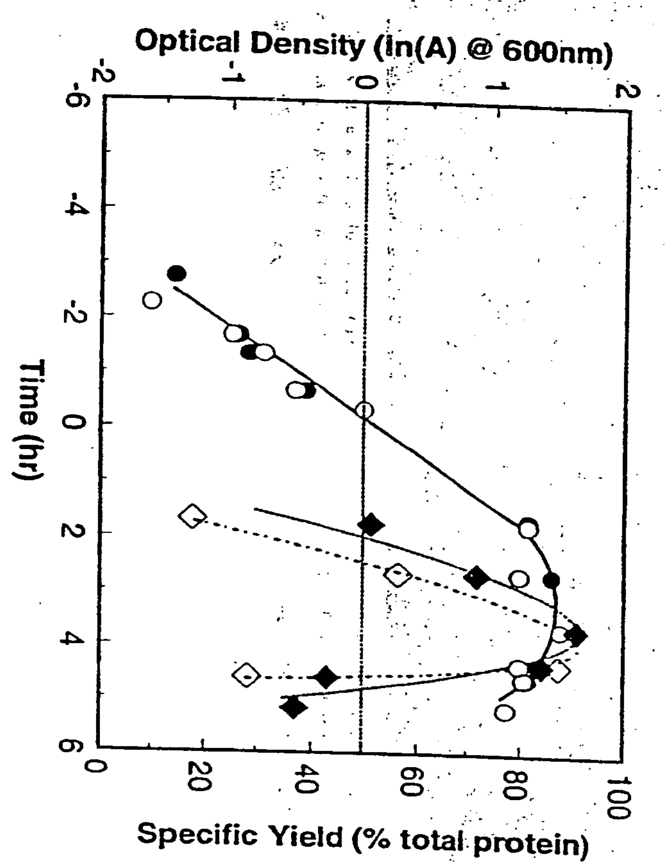

Optical Density (ln(A) @ 600nm)
-2 -1 0 1 2
Time (hr)
-6 -4 -2 0 2 4 6
Specific Yield (% total protein)
0 20 40 60 80 100

FIGURE 2

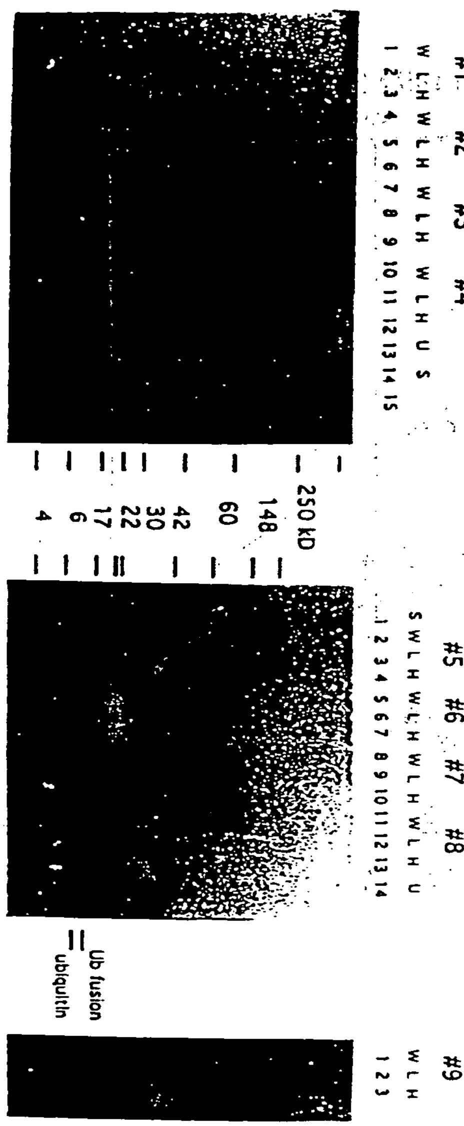

#1 #2 #3 #4
W L H W L H W L H W L H U S
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
250 kD
148
60
42
30
22
17
6
4
#5 #6 #7 #8
S W L H W L H W L H W L H U
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Ub fusion
ubiquitin
#9
W L H
1 2 3

FIGURE 3

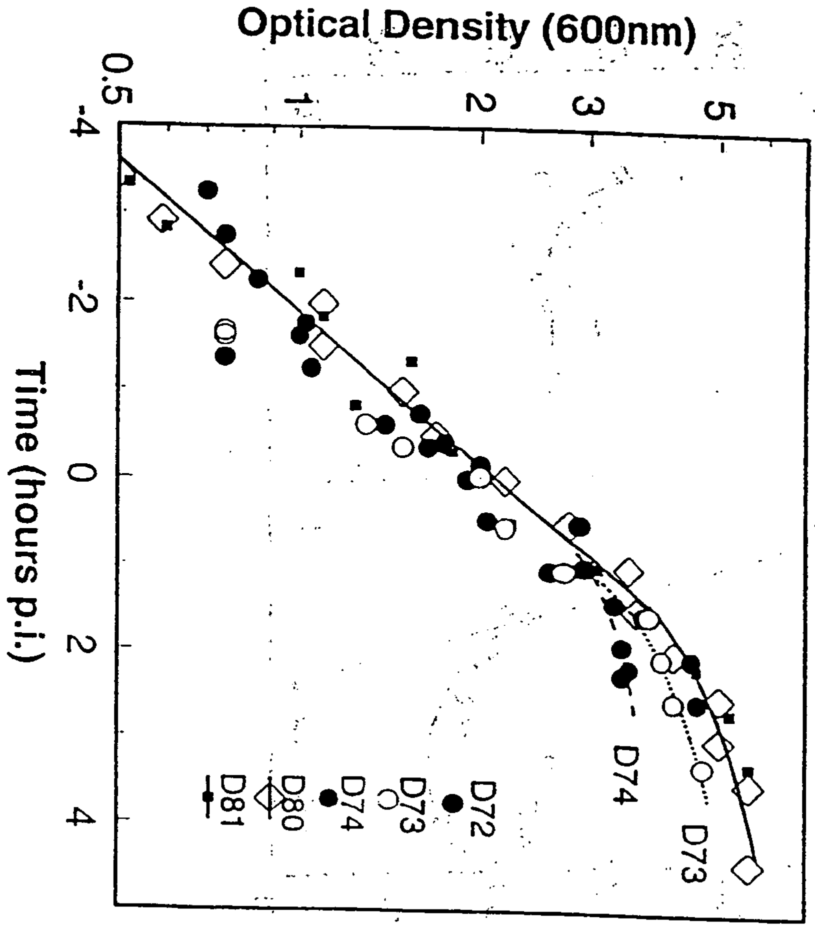

Optical Density (600nm)
0.5
1
2
3
5
Time (hours p.i.)
-4
-2
0
2
4
D74
D73
D72
D73
D74
D80
D81

FIGURE 4

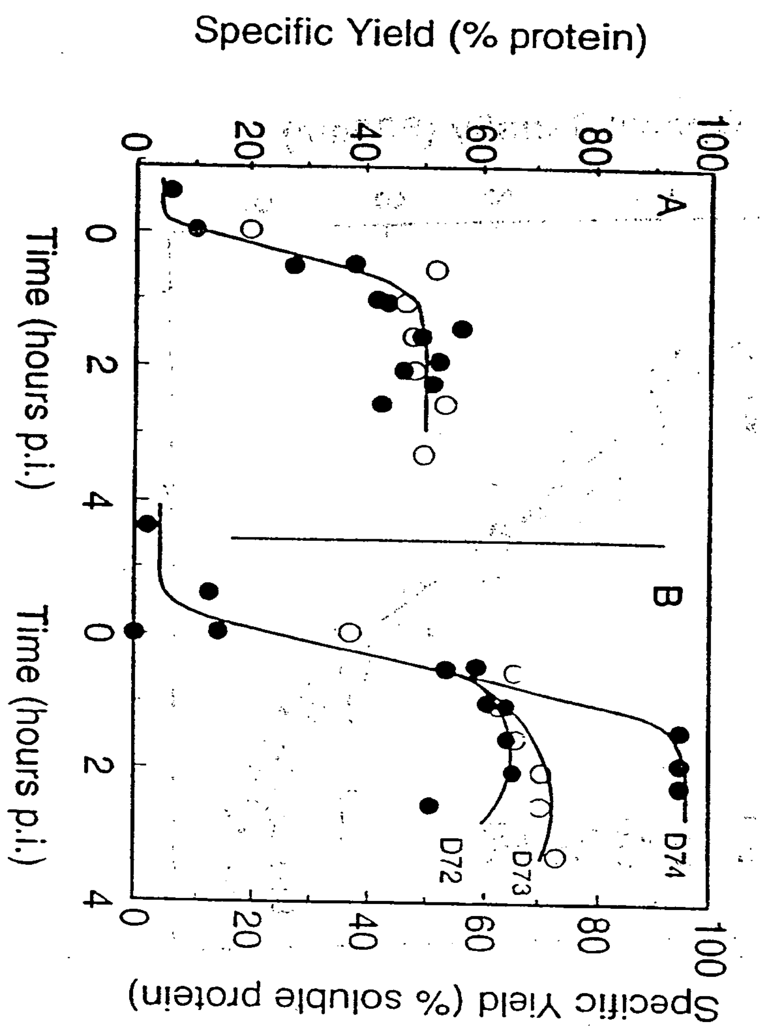

Specific Yield (% protein)
Specific Yield (% soluble protein)
Time (hours p.i.)
Time (hours p.i.)
A
B
D72
D73
D74

FIGURE 5

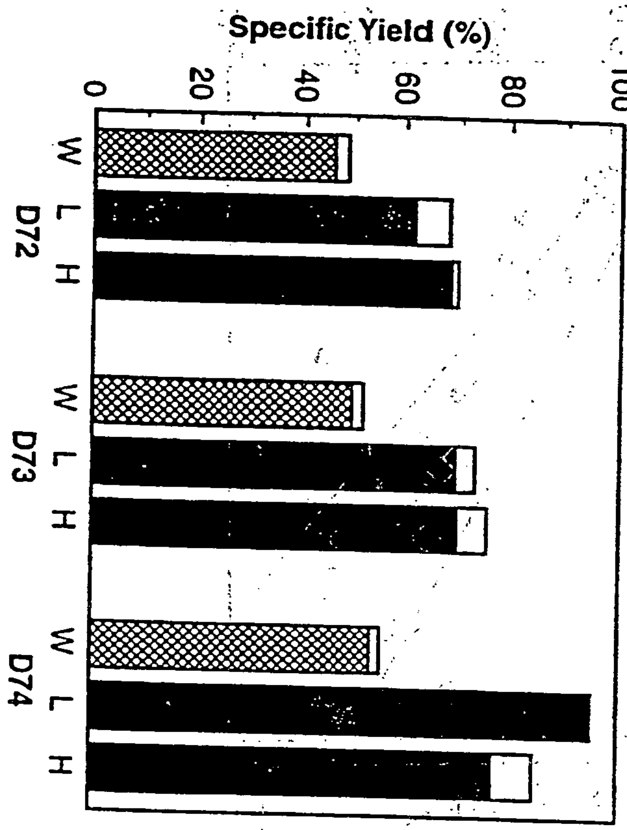


FIGURE 6

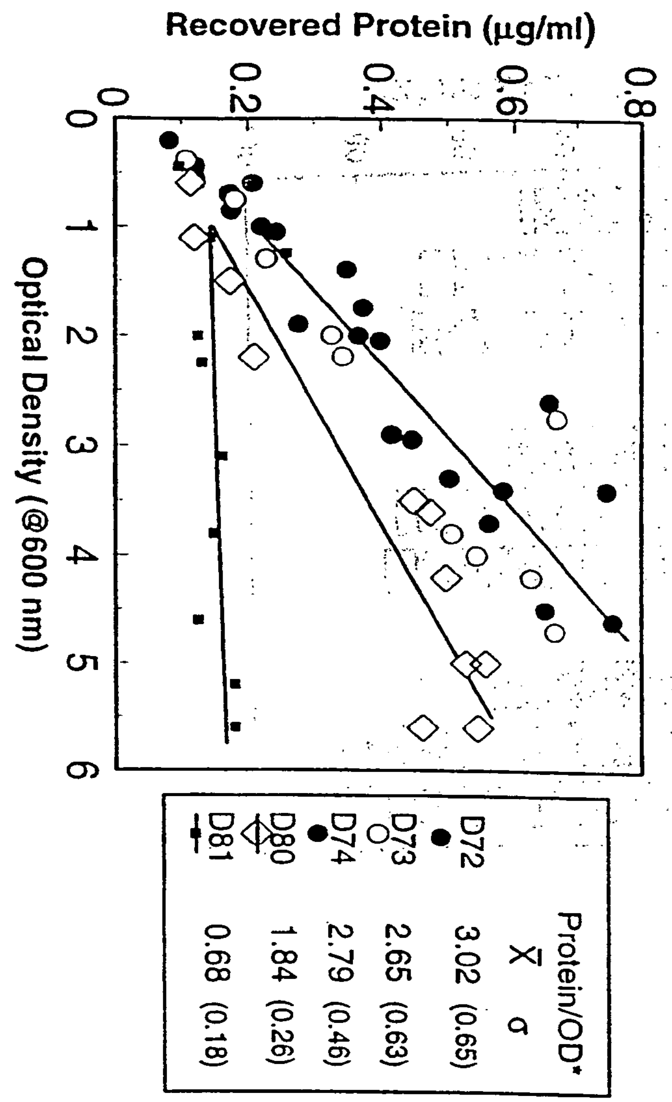
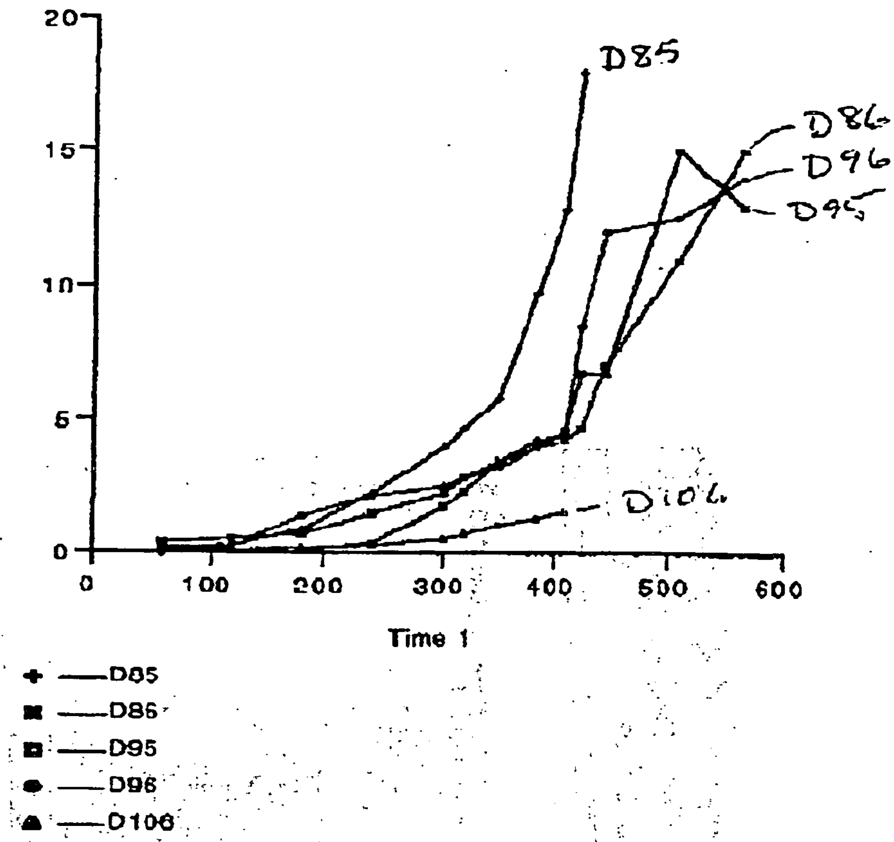

Recovered Protein (µg/ml)
0
0.2
0.4
0.6
0.8
Optical Density (@600 nm)
0
1
2
3
4
5
6
Protein/OD*
X̄ σ
D72 3.02 (0.65)
D73 2.65 (0.63)
D74 2.79 (0.46)
D80 1.84 (0.26)
D81 0.68 (0.18)

FIGURE 7



20
15
10
5
0
0
100
200
300
400
500
600
Time 1
D85
D86
D96
D95
D106
D85
D86
D95
D96
D106

FIGURE 8

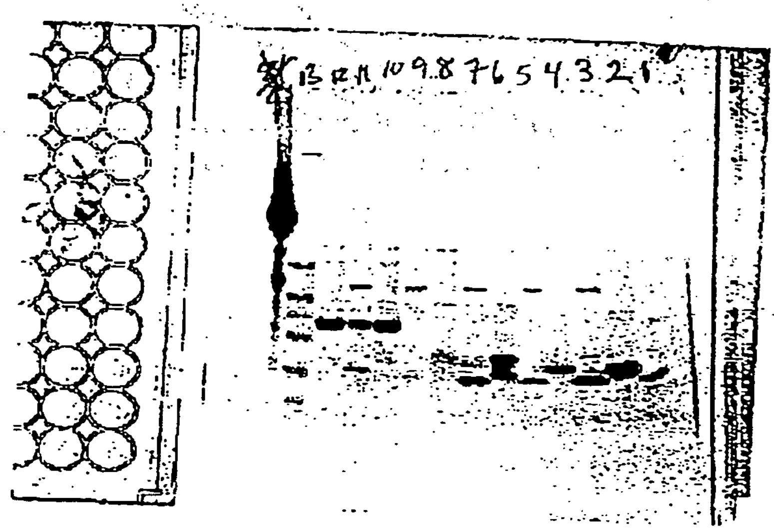


FIGURE 9

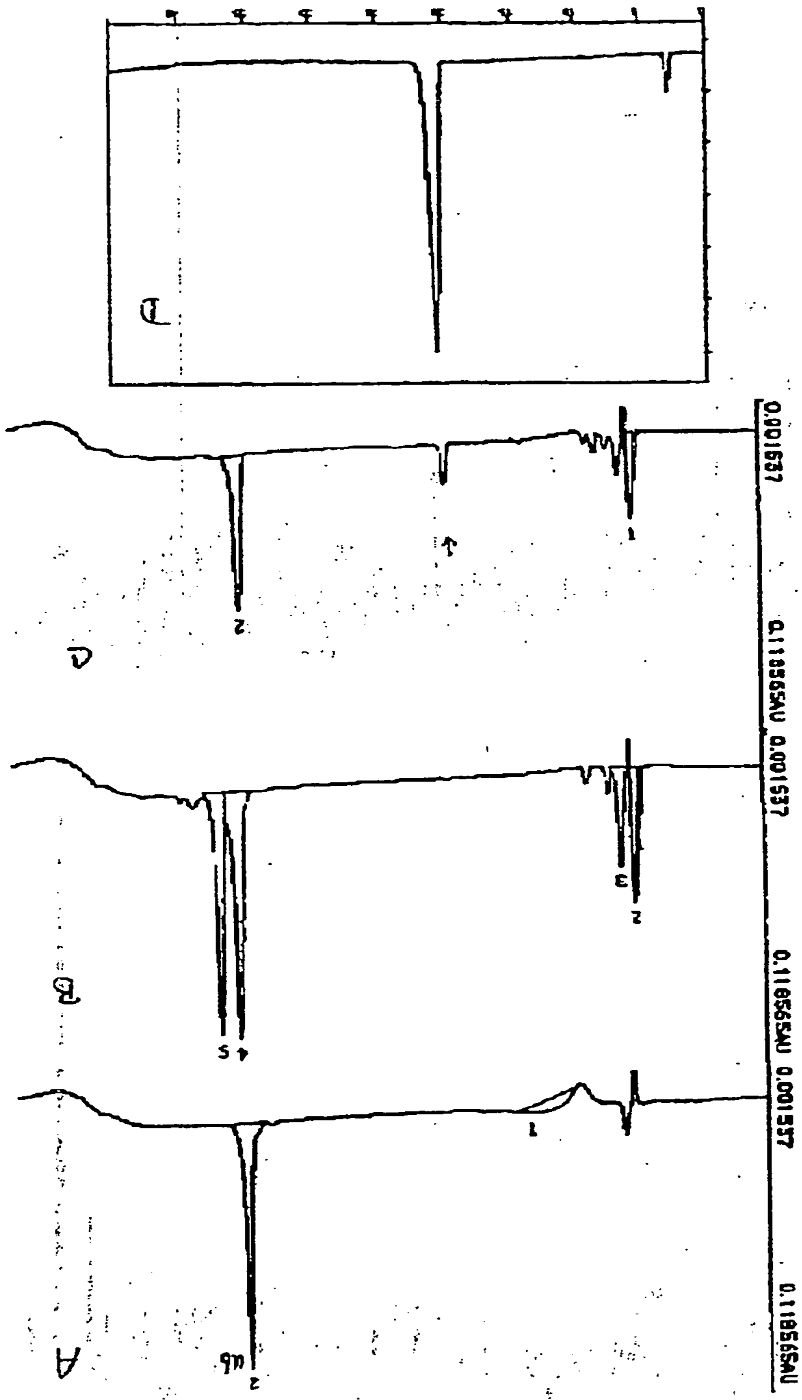


FIGURE 10

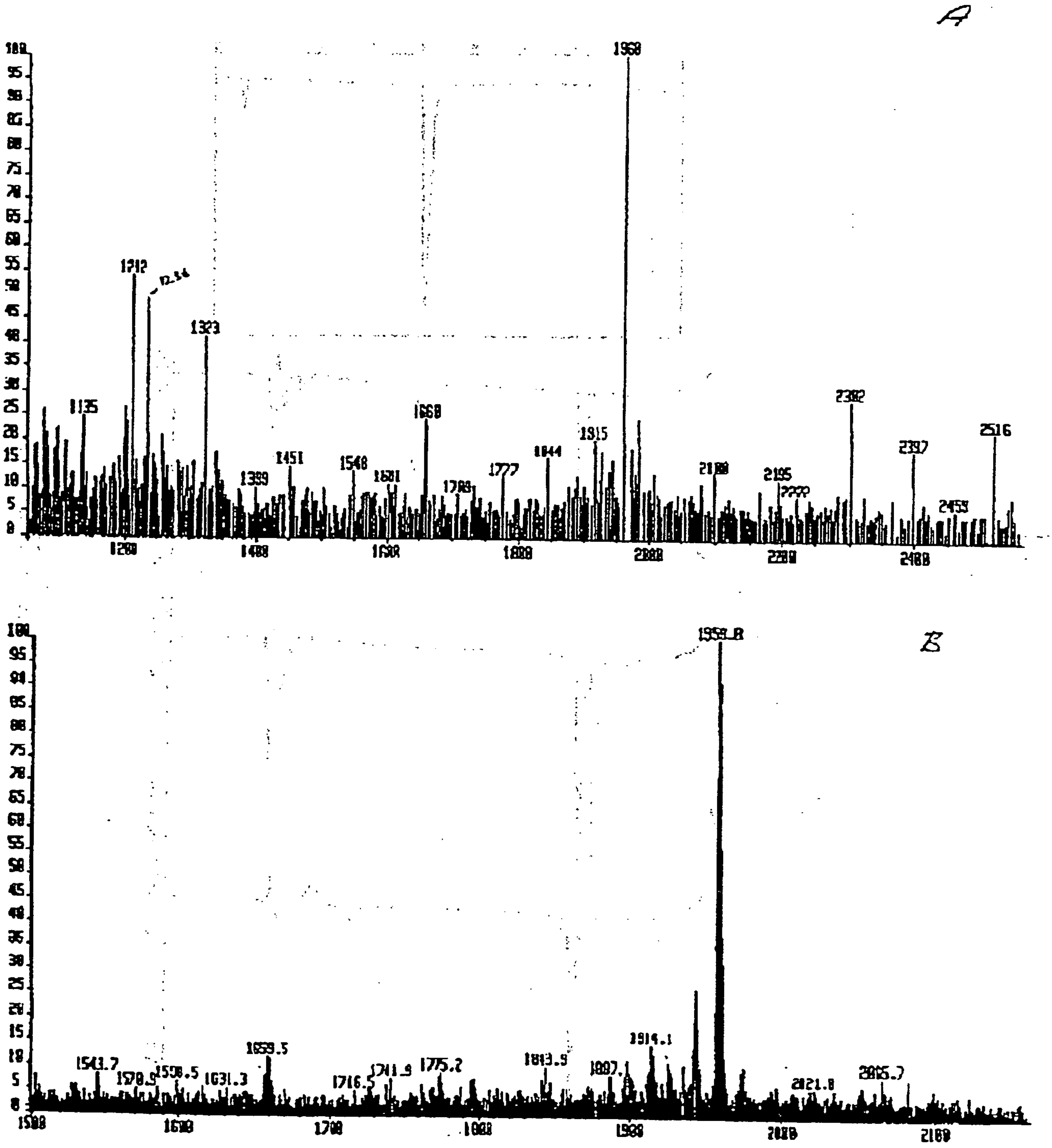
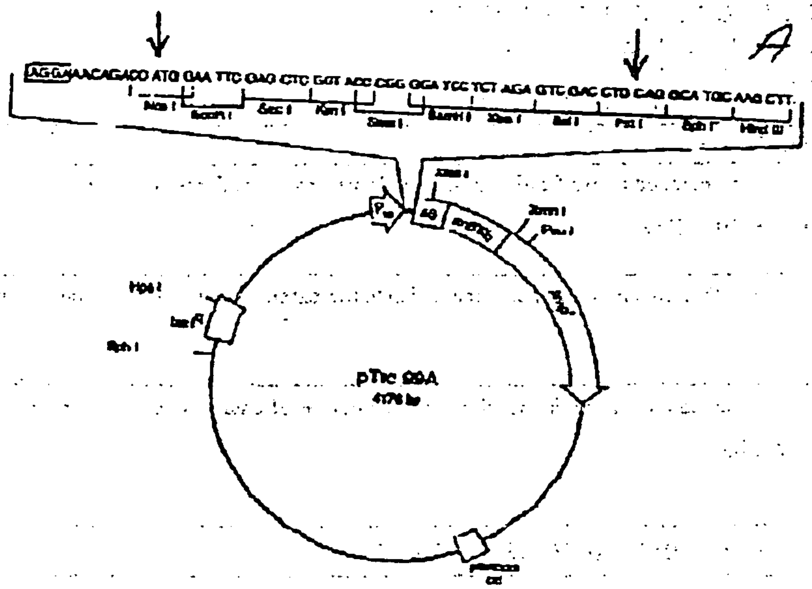

A
1960
1212
1323
1135
1660
1915
2302
2516
2397
1844
1451
1548
1399
1777
1708
2180
2195
2455
1200
1400
1600
1800
2000
2200
2400
B
1958.8
1659.5
1543.7
1631.3
1775.2
1914.1
1500
1600
1700
1800
1900
2000
2100

FIGURE 11

A



B

10	20	30	40	50	60
ATGCAGATCT	TCGTGAAAAC	CCTGACCGGC	AAAACCATCA	CTCTAGAAGT	GGAACCGAGC
70	80	90	100	110	120
GATACCATCG	AAAACGTGAA	AGCGAAAATC	CAGGATAAAG	AAGGCATCCC	GCCGGATCAG
130	140	150	160	170	180
CAGCGTCTGA	TCTTTGCGGG	CAAACAGCTG	GAAGATGGCC	GCACCCTGTC	TGATTATAAC
190	200	210	220	230	240
ATCCAGAAAG	AGTCGACCCT	GCATCTGGTC	TTAAGACTGC	GTGGCGGC..	....

FIGURE 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/10858

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
U.S. : Please See Extra Sheet.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LaBEAN et al. Design, expression, and characterization of random sequence polypeptides as fusions with ubiquitin. FASEB J. 01 April 1992, Vol. 6, page A471, see entire document.	1-40
Y	BUTT et al. Ubiquitin fusion augments the yield of cloned gene products in Escherichia coli. Proc. Natl. Acad. Sci. USA. 01 April 1989, Vol. 86, pages 2540-2544, see entire document.	1-40
Y	YOO et al. Synthesis of peptides as cloned ubiquitin extensions. J. Biol. Chem. 15 October 1989. Vol. 264, No. 29, pages 17078-17083, see entire document.	1-40

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 AUGUST 1996	02 OCT 1996

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	LISA J. HOBBS, PH.D.
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/10858

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,262,322 A (LIU ET AL.) 16 November 1993, column 2, lines 58-65 and column 3, lines 17-22.	1-40
Y,P	US 5,506,120 A (YAMAMOTO ET AL.) 09 April 1996, column 3, lines 54-64.	1-40

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/10858

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C12N 5/00, 1/20, 1/14, 1/16, 1/18; C12P 21/06; C07K 1/00

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/69.1, 240.1, 252.3, 252.33, 254.1, 272

B. FIELDS SEARCHED
Minimum documentation searched
Classification System: U.S.

435/69.1, 240.1, 252.3, 252.33, 254.1, 272

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, STN Bioscience Index (45 Databases Searched), ubiquitin#, fusion protein#, ferment?, A37, coli, yeast, insect, baculovirus, gram# per liter#, specific activity, cultur###